Sequence Context Dependence of Tandem Guanine:Adenine Mismatch Conformations in RNA: A Continuum Solvent Analysis

Gilberto Villescas-Diaz and Martin Zacharias*

Theoretische Biophysik, Institut für Molekulare Biotechnologie, Jena, Germany; and *School of Engineering and Science, International University Bremen, Germany

ABSTRACT Guanine:adenine (G:A) mismatches and in particular tandem G:A (tG:A) mismatches are frequently observed in biological RNA molecules and can serve as sites for tertiary interaction, metal binding and protein recognition. Depending on the surrounding sequence tG:A mismatches can adopt different basepairing topologies. In the sequence context (5′-) G**GA**C (tandem G:A in bold) a face-to-face (imino or Watson-Crick-like) pairing is preferred whereas in the C**GA**G context, G and A adopt a sheared arrangement. Systematic conformational searches with a generalized Born continuum model and molecular dynamics simulations including explicit water molecules and ions have been used to generate face-to-face and sheared tG:A mismatches in both C**GA**G and G**GA**C sequence contexts. Conformations from both approaches were evaluated using the same force field and a Poisson-Boltzmann continuum solvent model. Although the substate analysis predicted the sheared arrangement to be energetically preferred in both sequence contexts, a significantly greater preference of the sheared form was found for the C**GA**G context. In agreement with the experimental observation, the analysis of molecular dynamics trajectories indicated a preference of the sheared form in the case of the C**GA**G-context and a favorization of the face-to-face form in the case of the G**GA**C context. The computational studies allowed to identify energetic contributions that stabilize or destabilize the face-to-face and sheared tandem mismatch topologies. The calculated nonpolar solvation and Lennard-Jones packing interaction were found to stabilize the sheared topology independent of the sequence context. Electrostatic contributions are predicted to make the most significant contribution to the sequence context dependence on the structural preference of tG:A mismatches.

INTRODUCTION

Tandem guanine:adenine mismatches (5′GA3′/3′AG5′) are frequently observed in many biological RNAs, such as the hammerhead ribozyme (Pley et al., 1994), the Tetrahymena ribozyme (Cate et al., 1996a,b), and in ribosomal RNA (Ban et al., 2000; Schlünzen et al., 2000; Wimberley et al., 2000; Schlünzen et al., 2001), and can serve as sites for tertiary interactions as well as ligand binding (Gutell et al., 1994; Gautheret et al., 1994; Guzman et al., 1998). The mismatch order G:A followed by A:G is more prevalent in biological RNAs than the order A:G followed by G:A (Gutell et al., 1994; SantaLucia and Turner, 1993; Wu et al., 1997). Structural studies indicate that the conformation of the tandem G:A (tG:A) mismatch motif depends on the flanking sequences, and can adopt two main distinct basepairing topologies (SantaLucia and Turner, 1993; Wu and Turner, 1996; Heus et al., 1997; Zhang et al., 1998). In the G**GA**C context, a face-to-face (Watson-Crick type) imino proton pairing scheme is preferred, whereas in the C**GA**G context, G and A adopt a sheared pairing arrangement (SantaLucia and Turner, 1993; Wu and Turner, 1996). Both pairing schemes provide very different interaction surfaces accessible for proteins and other ligands in the RNA minor and major grooves (Fig. 1). In addition, the two alternative topologies deform the groove geometry to different degrees. The face-to-face arrangement widens and increases the accessibility of the RNA major groove (Wu and Turner, 1996). The sheared topology decreases the distance between the opposite RNA strands and creates an overall more compact structure. The overrepresentation of the sheared form in biological RNAs (Gautheret et al., 1994; Wu and Turner, 1996) has been attributed to the fact that only the sheared arrangement may allow for certain tertiary contacts involving base functional groups that are not accessible in the case of the face-to-face form. Based on an experimental estimate of the free energy necessary to switch from the face-to-face into a sheared arrangement of ~2–3 kcal mol^{-1} it has been speculated that protein-RNA binding may provide sufficient energy to promote such a switch, and in turn can lead to a global change in the RNA geometry (Wu and Turner, 1996). Understanding global conformational changes in RNA, and how they can be mediated by protein-RNA or RNA-RNA interactions, is of biological importance to understand the function of large RNA-containing biomolecules.

The experimental structure determination alone does not explain the structural preference for the two possible pairing geometries in tG:A mismatches in different sequence contexts. Aim of the present study is to use a Poisson-Boltzmann and a generalized Born continuum solvent model (Still et al., 1990; Hawkins et al., 1995, 1996; Jayaram et al., 1998; Srinivasan et al., 1998, 1999) to analyze the energetic con-

Submitted October 13, 2002, and accepted for publication February 27, 2003.

Address reprint requests to Martin Zacharias, School of Engineering and Science, International University Bremen, Campus Ring 1, D-28759 Bremen, Germany. Tel.: 49-421-200-3541; Fax: 49-421-200-3249; E-mail: m.zacharias@iu-bremen.de.

Abbreviations used: EM, energy minimization; FDPB, finite difference Poisson-Boltzmann; ftf, face to face; GB, generalized Born; MD, molecular dynamics; NMR, nuclear magnetic resonance; MM/PBSA, molecular mechanics/Poisson-Boltzmann surface area; tG:A, tandem guanine:adenine.

0006-3495/03/07/416/10 $2.00

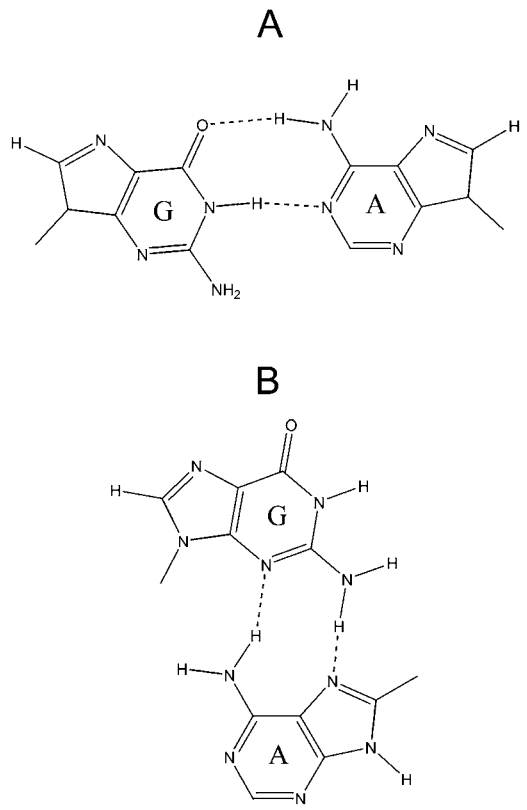

FIGURE 1 Tandem G:A mismatches in the face-to-face (*A*) and sheared (*B*) basepairing forms.

tributions that stabilize or destabilize the two possible tG:A basepair topologies in RNA and how neighboring basepairs affect their relative stability. Such a simulation model is only an approximation to reality. However, it has already been used successfully to study conformational preferences of single base bulges and hairpin loop structures in RNA and DNA, in good agreement with experimental results (Srinivasan et al., 1998; Zacharias and Sklenar, 1999; Zacharias, 2001). Generalized Born-type solvation models have also been used to study the dynamics of nucleic acids (Williams and Hall, 1999, 2000a,b; Tsui and Case, 2000) in very reasonable agreement with simulations in the presence of explicit solvent (Tsui and Case, 2000).

Systematic conformational searches have been employed in the current study to investigate structural preferences of tG:A mismatches. In this approach, only energy-minimized conformational substates have been considered. In a second approach, ensembles of tG:A conformations obtained from a molecular dynamics simulation in the presence of explicit solvent and ions have been analyzed using the same continuum solvent model and force field as used to evaluate conformations from the conformational search. The results of both approaches show qualitative agreement and are compatible with the experimentally observed conformational preference of tG:A mismatches. For the sheared tandem tG:A arrangement, two low energy subtopologies were found in both the substate analysis as well as the molecular dynamics simulations that differ in the hydrogen-bonding pattern, and that are also observed in experimental G:A mismatch-containing structures. The analysis of energetic contributions allows to elucidate the origin of the sequence context-dependence of the structural preference and offers some insight into interactions that stabilize the sheared and face-to-face tG:A topologies.

METHODS

A modified version of the Junction Minimization of Nucleic Acids (JUMNA) program (Lavery et al., 1995) in combination with the Cornell et al. (1995) force field was used for energy minimization (EM) and conformational search. The energy function consisted of pairwise additive nonbonded Coulomb and Lennard-Jones terms (no cutoff) and valence angle and dihedral angle contributions. During EM electrostatic reaction field contributions (E_{reGB}) due to differences in the dielectric constants assigned to the RNA molecule ($\varepsilon_{in} = 1.0$) and surrounding aqueous solvent ($\varepsilon_w = 78.0$) were calculated using the generalized Born (GB) model (Still et al., 1990). The α_{ij} were calculated using the pairwise descreening approximation described by Hawkins et al. (1995, 1996). The modified Bondi set of atomic radii (Bondi, 1964) as derived by Tsui and Case (2000) was used with the following values (in Å): $R_H = 1.3$ (aliphatic hydrogens), $R_{H-N} = 1.2$ (for hydrogens connected to nitrogen atoms), $R_H = 0.8$ (for hydrogens connected to sugar O2′ or O3′ atoms), $R_C = 1.7$, $R_N = 1.55$, $R_O = 1.5$, and $R_P = 1.85$. The descreening parameters were $S_H = 0.85$, $S_C = 0.72$, $S_N = 0.79$, $S_O = 0.85$, and $S_P = 0.86$, with a radius offset parameter of -0.125 Å. For the final minimized structures, electrostatic reaction field and salt contributions were also calculated with the finite-difference Poisson-Boltzmann (FDPB) method as implemented in the University of Houston Brownian Dynamics program (Madura et al., 1995), and the same atom radii, as used in the GB model. For calculations in the presence of salt, the nonlinear FDPB was solved (Sharp and Honig, 1990).

Surface area-dependent nonpolar solvation contributions (ΔE_{SASA}) were evaluated from the accessible surface area (Shrake and Rupley, 1973), with a surface area tension coefficient of $\gamma = 0.00542$ kcal $\times$ mol^{-1} Å^{-2} (Sitkoff et al., 1994). This term was only calculated for the final energy-minimized structures since it varies very little between different conformers of one topology ($<\sim 0.1$ kcal $\times$ mol^{-1}). The total energy of a conformer (ΔE_{totPB} or ΔE_{totGB}) is given as a sum of Coulomb (ΔE_{Coul}), Lennard-Jones (ΔE_{LJ}), valence and torsion angle (ΔE_{TA}), electrostatic solvation (ΔE_{reGB}, ΔE_{rePB} or $\Delta E_{rePBsalt}$), and nonpolar solvation contributions (ΔE_{SASA}).

RNA structures and conformational search

The experimentally determined RNA structures for the two self-complementary sequences (*A*: (rGCG**GA**CGC)$_2$) and (*B*: (rGGC**GA**GCC)$_2$) were used as start structures for conformational searches. The experimental structure for sequence *A* indicates a face-to-face G:A pairing (Wu and Turner, 1996; see Fig. 1). Based on its structure the JUMNA program was used to generate a face-to-face model for sequence *B*. Vice versa the structure for sequence *B* with an experimentally observed sheared G:A tandem mismatch (SantaLucia and Turner, 1993) served as the template for a sheared model structure for sequence *A*. Systematic conformational searches were performed starting from the four structures. During the conformational search, various combinations of backbone torsion-angle window constraints were applied during constraint energy minimization for the mismatch nucleotides and adjacent nucleotides (other nucleotides were kept in *A*-form). For the three central dinucleotide steps, the torsion angles α were constraint to two windows ($-70°$; $-50°$) or (60°; 120°), γ to (50°; 70°) or (170°; $-170°$), and ζ to ($-70°$; $-50°$) or (170°; $-170°$), and the δ-torsion angles of the two central nucleotides (G and A) were constraint to either the

C3′-endo (70°; 90°) or C2′-endo (140°; 160°) regimes, respectively, resulting in 2048 combinations (only symmetric constraints on both strands have been considered). All structures were subsequently relaxed by unconstraint energy minimization. For a subset of low energy conformers, the orientation of the central 2′-OH groups was systematically scanned to identify an optimal orientation. The 30 lowest energy structures from each search were collected and evaluated using the FDPB approach.

Molecular dynamics simulations

The lowest energy structures obtained from the four conformational searches served as start structures for molecular dynamics (MD) simulations. The MD simulations were performed with the *sander* module of the AMBER5 package (Pearlman et al., 1995) in a periodic box including TIP3 water (Jorgensen et al., 1983) and ions, with the PME method to account for long-range electrostatics in a periodic system and using the Cornell et al. (1995) force field. Initial positions of sodium counterions and five additional sodium and chloride counterions were placed using the *xleap* module of the AMBER package. Approximately 2000 water molecules were added to fill the box. A 9-Å cutoff for the short-range nonbonded interactions was used in combination with the particle mesh Ewald option, using a grid spacing of ~0.9 Å to account for long-range electrostatic interactions. After a heating and equilibration phase of the solvent and ions and subsequent heating of the whole system in 50-K steps (up to 300 K) within 0.1 ns, each system was equilibrated for 1.0 ns at 300 K followed by 3 ns data-gathering time. A set of 200 solute structures along the trajectories were postprocessed using the FDPB continuum solvent model described above (molecular mechanics Poisson-Boltzmann surface area—MM/PBSA—method; see Srinivasan et al., 1998, and Kollman et al., 2000). For each solute structure in the trajectory file, the same molecular mechanical energy as for the above energy-minimized conformations consisting of bonded and nonbonded energy terms of the Cornell et al. (1995) force field was calculated, supplemented with a polar and nonpolar solvation contribution. The solvation contributions were calculated in the same way and with the same parameters as used for the evaluation of conformers generated during the conformational search. Cartesian coordinate fluctuations for all heavy atoms were calculated after subtraction of overall translation and rotation.

RESULTS

Energy minimization and conformational search

Conformational searches involving 2048 different backbone torsion angle combinations for the central face-to-face or sheared tG:A mismatch motif and flanking basepairs were performed for the two sequence contexts 5′-GGC**GA**GCC (in the following: C**GA**G) and 5′-GCG**GA**CGC (in the following: G**GA**C), respectively. Energy minimization of each conformer was performed including a GB continuum solvent model for the electrostatic interactions. For comparison in each of the four cases, a subset of low energy conformations was also evaluated using the FDPB approach. A reasonable correlation between electrostatic reaction field energies for both approaches was found. The correlation was better for the face-to-face form. Therefore, the relative ranking of low energy substates correlates quite well for both electrostatic models in the case of the ftf-tG:A-conformations but less well in the case of the sheared tG:A arrangement (see Table 1). Since the FDPB model is principally more accurate than the GB electrostatic model, it served as the reference model for the analysis of the results (including 0.15 M salt and solving the nonlinear PB).

The calculated lowest energy structures obtained from the conformational search are in good agreement with available experimental results. In the case of the G**GA**C sequence context an NMR structure is known for the ftf-tG:A arrangement (Wu and Turner, 1996). The Root-mean-square deviation (RMSD) between the lowest energy structure from the search and the experimental (average NMR) structure is <1.5 Å (Fig. 2). A characteristic feature of the experimental structure is an α-γ flip (a concerted change of the backbone

TABLE 1 Ranking of tandem G:A mismatch conformers from conformational searches

	$\Delta E_{\text{totPBsalt}}$	ΔE_{totPB}	ΔE_{totGB}	ΔE_{rePB}	ΔE_{reGB}	ΔE_{Coul}	ΔE_{Elec}	ΔE_{LJ}	ΔE_{SASA}	ΔE_{TA}
C**GA**G (Sheared)										
1 ε/ζ(4) type I	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2 α/γ(5)	3.4	0.6	2.4	135.9	137.6	−144.1	−8.2	4.6	0.6	3.6
3 ε/ζ(4) type II	2.9	3.4	−2.7	−22.1	−28.3	24.3	2.2	−1.7	−0.3	3.2
4 ε/ζ(4,5) α/γ(6)	9.3	9.3	3.0	39.6	33.2	−38.4	1.2	5.6	−0.3	2.9
5 ε/ζ(4) α/γ(4)	10.7	10.0	2.8	62.0	54.7	−63.2	−1.2	5.8	0.0	5.3
C**GA**G(ftf)										
1 α/γ(4)	7.6	7.0	7.3	106.4	106.7	−108.4	−2.0	10.7	0.8	−2.7
2 A-form	7.3	7.6	7.2	32.3	32.5	−32.1	−0.2	11.5	0.8	−4.4
C**GA**G (Sheared)										
1 ε/ζ(4) type I	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2 α/γ(5)	2.6	0.3	1.5	107.4	108.9	−118.5	−11.1	6.4	0.5	4.3
3 ε/ζ(4) type II	2.7	3.6	−3.3	−19.8	−26.6	21.5	1.7	−0.6	−0.4	2.7
4 ε/ζ(4,5)α/γ(6)	5.5	6.0	6.4	33.9	28.7	−37.2	−3.3	8.5	−0.2	0.9
5 ε/ζ(4) α/γ(4)	12.0	11.4	1.5	54.2	44.6	−52.0	2.2	4.8	−0.2	4.4
G**GA**C(ftf)										
1 α/γ(4)	3.8	3.6	3.6	100.4	100.5	−105.5	−5.0	11.9	0.8	−4.2
2 A-form	5.1	6.4	3.9	29.6	27.3	−30.4	−0.8	12.6	0.8	−5.8

Energetic contributions are given in kcal mol^{-1} and with respect to the lowest energy sheared tG:A substate (this state shows best agreement with the experimental sheared tG:A structure; SantaLucia and Turner, 1993). The main deviations with respect to A-form backbone structure (nucleotide positions are given in parentheses) of each substate are indicated in the first column. The second column corresponds to the total energy of the conformer with respect to the FDPB approach (+0.15 M salt). For the other energy terms, see Methods section.

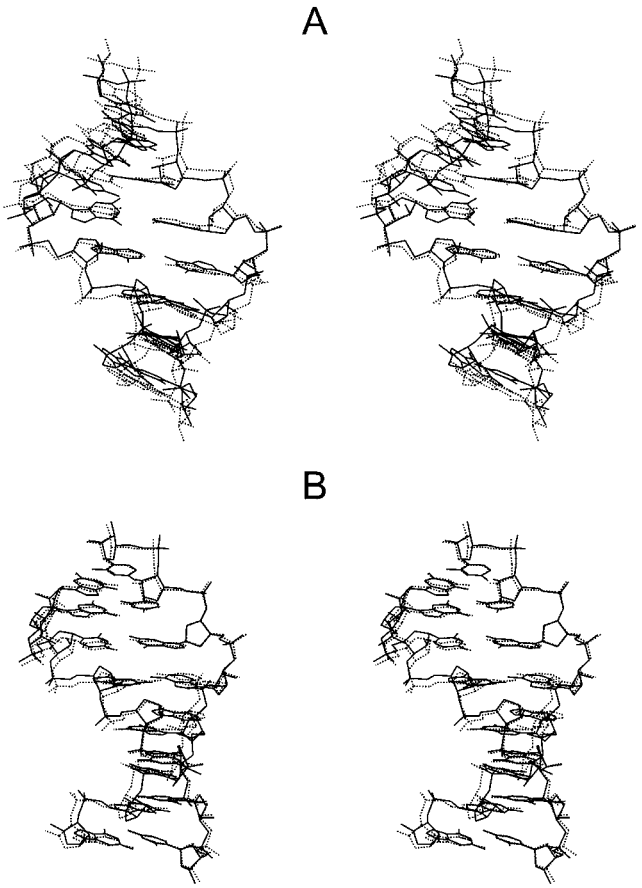

FIGURE 2 Comparison of tandem G:A mismatch structures in stereo in the face-to-face arrangement (*A*, sequence: (rGCG**GA**CGC)$_2$) and sheared topology (*B*, sequence: (rGGC**GA**GCC)$_2$). The low energy structures obtained from the conformational searches (*bold line*) that showed best agreement with experiment are superimposed on corresponding experimental NMR-derived structures (*dashed lines*, *A*: pdb-entry 1mis, Wu and Turner, 1996; *B*: pdb-entry 1yfv, SantaLucia and Turner, 1993), respectively.

torsion angle α and γ from –gauche and +gauche, respectively, toward a trans state) at the central G-A step. This same backbone feature was also found in the lowest energy ftf-tG:A-structure from the conformational search (see Tables 1 and 2). An α-γ flip can cause a slight increase in the distance between opposite strands compared to regular RNA allowing for more space to accommodate a large G:A pair compared to a regular Watson-Crick basepair. However, the second-lowest energy substate showed a backbone conformation typical for A-form RNA. Other low energy structures (not included in Table 1) differ from the lowest energy structure in additional α-γ flips at neighboring nucleotides and in different orientation of the 2′-OH group. In the case of the C**GA**G sequence context, the structure with an α-γ flip at the central A-G step was found to be almost isoenergetic with the fully A-form backbone geometry (Table 1). The optimal orientation of the 2′-OH group for each low energy substate was investigated separately. Both for the A-form flanking nucleotides as well as for the central tG:A (ftf) mismatches, three minimum energy orientations were found. These minima correspond to H_2'-C_2'-O_2'-HO dihedral angles of ~80°, 190°, and 310°, respectively, and closely match the orientations found in explicit solvent simulations of RNA (Auffinger and Westhof, 1997). Of these minima the orientation that points toward the sugar O3′-atom (H_2'-C_2'-O_2'-HO dihedral angles of ~80°) was generally the lowest energy orientation. This orientation is also the preferred orientation found in MD simulations of RNA (Auffinger and Westhof, 1997). It should be noted that this is not a trivial result, since simpler electrostatic models do not necessarily agree with the optimal 2′-OH orientation found in MD simulations. In Table 1, only conformers with optimal 2′-OH orientations have been listed.

In the case of the sheared tG:A mismatches, the conformer ranking between the two electrostatic models showed larger differences. The lowest energy conformer (using the PB model with added salt as reference) from both approaches, and in both sequence contexts, showed the best agreement with available experimental results. An NMR-derived structure has been determined for the C**GA**G context in the sheared tG:A conformation (SantaLucia and Turner, 1993). The RMSD between average NMR structure and the lowest energy conformer was <1.4 Å (Fig. 2). In addition, a characteristic feature of the backbone topology found in the experimental structure that is a B_{II} state characterized by an increase of the ε-torsion angle and decrease of the ζ-angle at the central step was also found in the calculated lowest energy substate (Table 2). The B_{II} state at the central dinucleotide step was found in several low energy substates. A B_{II} state can promote a slight decrease in the distance between opposite strands and therefore supports a sheared G:A pairing that requires a reduced distance between opposite strands. Additional backbone variations include α-γ flips at dinucleotide steps adjacent to the central step. For the 2′-OH group at the central dinucleotide step of the low energy sheared structures, only one optimal orientation (minimum) was found (the 2′-OH vector pointing approximately toward the O3′ atom). Two subforms of G:A sheared basepairing arrangement were found (type I and II). These two subforms differ slightly in the distance and relative orientation of the guanine and adenine bases in the mismatch and hydrogen-bonding geometry (Fig. 3). In the type I two strong (short distance) hydrogen bonds are formed between the two G:A pairs. This form corresponds to the geometry found in the NMR structure of tG:A mismatches in the C**GA**G sequence context. In the subform II, only one short distance hydrogen bond is formed in the mismatches; however, in this case, a hydrogen bond between the N_6H_1 of the central adenine and the O_1P of the central dinucleotide step was observed. This G:A mismatch geometry has also been found experimentally; for example, as a closing pair in the x-ray structure of GNRA tetraloops (see Fig. 3). The GB approach appears to generally slightly favor type II, whereas the FDPB

TABLE 2 Comparison of experimental and calculated lowest energy backbone structures

	α (P-O5′)	β (O5′-C5′)	γ (C5′-C4′)	δ (C4′-C3′)	ε (C3′-O3′)	ζ (O3′-P)
Sheared tandem G:A mismatch						
C	−73.6 (−77.9)	173.8 (174.4)	62.5 (59.3)	79.3 (88.7)	−155.4 (−158.9)	−50.5 (−53.6)
G	−77.1 (−81.1)	178.4 (−177.8)	54.4 (54.4)	82.6 (90.6)	−177.1 (−179.0)	−107.4 (−106.1)
A	−66.1 (−73.5)	173.5 (−165.4)	62.6 (55.7)	81.5 (94.4)	−146.9 (−138.5)	−48.6 (−56.4)
G	−74.6 (−67.6)	172.1 (165.0)	61.1 (61.6)	83.1 (91.4)	−156.3 (−156.7)	−63.5 (−67.6)
Face-to-face tandem G:A mismatch						
G	−75.3 (−73.8)	−179.9 (179.7)	56.8 (61.0)	80.0 (79.9)	−160.4 (−167.7)	−76.6 (−72.8)
G	−76.5 (−74.8)	175.7 (179.4)	59.6 (62.5)	82.2 (78.5)	−171.4 (−171.8)	−69.5 (−56.9)
A	143.8 (141.5)	−177.3 (−169.4)	−175.0 (−176.4)	85.5 (85.5)	−146.1 (−149.8)	−56.5 (−60.0)
C	−71.4 (−72.0)	−168.7 (−175.8)	55.5 (55.3)	78.6 (78.1)	−156.5 (−164.5)	−69.4 (−63.2)

Torsion angles are given for the lowest energy conformations found for the 5′-GGC**GA**GCC (sheared form) and 5′-GCG**GA**CGC (ftf form) sequences, respectively. Available torsion angles for the corresponding experimental structures (sheared form, SantaLucia and Turner, 1993; and ftf form, Wu and Turner, 1996) are given in parenthesis.

approach predicts a lower calculated energy for conformers with type I sheared G:A geometry.

For both sequence contexts, the energy difference between lowest energy sheared and ftf tG:A mismatch structures was smaller than 8 kcal mol^{-1}. This is due to a balance between various contributions that favor or disfavor one or the other form. A general trend for both sequence contexts is that the sheared arrangement is favored by van der Waals packing interactions, the surface area-dependent nonpolar solvation term, and slightly disfavored by more positive torsion and valence angle contributions (Table 1).

Comparison of equivalent low energy substates in the sheared tG:A conformation for the two sequence contexts with any selected low energy substate from the ftf-tG:A set of conformations (or vice versa) reveals a stronger favorization of the sheared over the ftf form in the case of the C**GA**G compared to the G**GA**C sequence context. In the case of comparing the lowest energy substates that also show best agreement with experimentally available tG:A structures, the calculated favorization is ~−7.6 kcal mol^{-1} in the C**GA**G context. For the G**GA**C sequence, the calculated energy difference between the corresponding lowest energy ftf and sheared tG:A-forms is ~−3.8 kcal mol^{-1}. The greater favorization of the sheared form in the C**GA**G case compared to the G**GA**C sequence context agrees with experiment. However, the calculation also predicts a lower energy of the sheared vs. face-to-face form in the G**GA**C context. This seems to contradict the experimental observation. It is important to note that conformational entropy contributions due to differences in the flexibility of the two tG:A mismatch forms are not included in the continuum solvent analysis. Results from MD simulations (see below) indicate a greater conformational flexibility of the face-to-face conformation, which makes a favorable entropic contribution to the stability of the face-to-face pairing geometry. Depending on the selected pair of substates, electrostatic contributions can either stabilize or destabilize the ftf form compared to the sheared substate. However, for a given pair of substates, the magnitude of the electrostatic energy difference tends to be larger for the G**GA**C than the C**GA**G context. Considering the calculated lowest energy substates that showed best agreement with available experimental structures (substate 1 for sheared and ftf forms in Table 1) as the dominant structural states, the electrostatic interactions overall stabilize the ftf form by 5 kcal mol^{-1} in the G**GA**C context, but by only 2 kcal mol^{-1} in the C**GA**G context.

For the comparison of lowest energy substates, the van der Waals and Coulomb contributions were further split into intrachain and interchain contributions (Table 3). Note that this is not possible for the reaction field term, which depends on shape and charge distribution of the entire molecule. The pattern of intra- and interstrand van der Waals contributions to the energy difference between ftf and sheared forms is relatively similar for both sequence contexts. Most of the van der Waals energy difference between sheared and face-to-face forms is due to interstrand (cross-stacking) contributions (partially compensated by intrastrand interactions) that significantly stabilize the sheared form. The interstrand base-base van der Waals interactions appear to favor the sheared form slightly more in the G**GA**C context than in the C**GA**G context. The Coulomb energy difference between the two forms overall strongly favors the ftf topology in both sequence contexts (see also Table 1). Interestingly, the calculated total interstrand Coulomb interaction more significantly favors the ftf form in the case of the G**GA**C context (−82.3 kcal mol^{-1}) than in the C**GA**G context (−47.2 kcal mol^{-1}). Note that these calculated total interaction energy differences include contributions not only from all nucleobases but also from the sugar-phosphate backbone interactions. The base-base charge interactions strongly stabilize the face-to-face form in the G**GA**C context (−10.5 kcal mol^{-1}) whereas a calculated destabilization is found in the C**GA**C context (7.8 kcal mol^{-1}). These results indicate that electrostatic base-base complementarity makes a significant contribution to the sequence context-dependence of the tG:A topology.

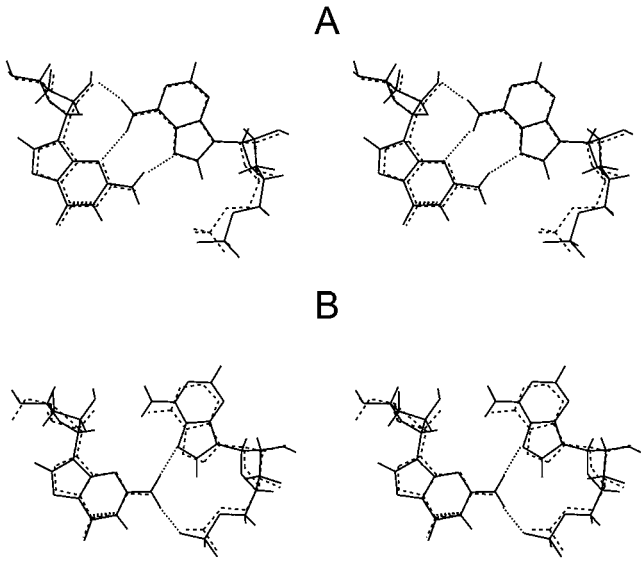

FIGURE 3 Comparison of two low energy subtopologies of the tandem G:A-mismatches in the sheared arrangement. The arrangement in form I (*continuous line*, hydrogens included) is characterized by three strong (small distance) hydrogen bonds between adenine (*strand 1*) N_7 and guanine (*strand 2*) H_2N_2, between adenine H_2N_6 and guanine N_3, and between adenine H_1N_6 and the O'_2 of guanine (indicated as *short dashed lines*). This topology is similar to the topology found experimentally in the NMR structure by SantaLucia and Turner (1993) and for example also in the x-ray structure of the hammerhead ribozyme (pdb-entry 1hmh, Pley et al., 1994). The G:A basepair formed by nucleotides G120 and A90, respectively, of the hammerhead ribozyme structure is superimposed (*dashed line* in *A*, no hydrogens included) on the calculated type I basepair. In the alternative low energy sheared G:A basepairing arrangement (type II, *continuous line* in *B*, hydrogens included) two short distance hydrogen bonds (between N_7 of adenine and H_2N_2 of guanine and O_1P of adenine and H_1N_2 of guanine) are formed (*short dashed lines* in *B*). A G:A sheared basepair similar to subform II is also present as a closing basepair of the GAAA tetraloop in the hammerhead x-ray structure (nucleotides G21 and A24 of pdb-entry 1hmh, *dashed lines* in *B*, no hydrogens). In *A* and *B*, the guanine corresponds to the left base of the basepairs.

Molecular dynamics simulations of tandem GA mismatch structures

Molecular dynamics simulations of up to 4 ns were performed starting from the lowest energy ftf-tG:A (second lowest energy in the C**GA**G context) and sheared tG:A mismatch structures (the conformations that showed the best agreement with available experimental structures). In all four simulations the tG:A conformations stayed reasonably close to the corresponding start structures and showed reasonable convergence of the RMSD from the start structure after ~1 ns (Fig. 4). No rearrangements of the initial tG:A topologies such as a transition from a ftf to a sheared topology or vice versa have been observed during the MD simulations. Nevertheless, for the sheared forms, transitions between type I and II tG:A arrangements (see Fig. 3) have been found during the simulation (with the type I arrangement more frequently sampled than type II, not shown). For both sequence contexts overall, slightly larger Cartesian coordinate fluctuations were observed in the case of the ftf-tG:A simulations (Fig. 5). To some degree, this is expected, since the sheared tG:A mismatches form an overall more compact structure. This result indicates a larger conformational freedom or conformational entropy of the ftf form compared to the sheared topology. Trajectory frames from the 2- to 4-ns interval of each MD simulation were analyzed with the FDPB continuum solvent model (Fig. 6). The continuum solvent analysis was performed with the same force field model and set of parameters as used to evaluate energy-minima from the conformational searches.

The continuum solvent analysis of the four trajectories showed qualitative agreement with the conclusions obtained from the comparison of low energy tG:A substates (see above). In the case of the 5′-C**GA**G sequence context, the calculations indicate an energetic preference of the sheared form compared to the ftf form of -10.7 kcal mol^{-1} (Table 4). In contrast, for the G**GA**C sequence, the imino-paired ftf form was favored by ~13 kcal mol^{-1} over the sheared tG:A mismatch arrangement. With an internal dielectric constant, $\varepsilon_{in} = 2$, a smaller calculated preference was obtained. In this case the calculated preference for the sheared form in the C**GA**G sequence context was ~ -8.0 kcal mol^{-1} and a preference for the ftf form in the case of the G**GA**C sequence of ~ -4.0 kcal mol^{-1} was obtained. The result is in qualitative agreement with the experimental observation. However, the calculated energetic preference for each topology is larger

TABLE 3 Comparison of intra- and interstrand contributions to the calculated tG:A mismatch stability

Ftf vs. sheared tG:A	5′-CGAG		5′-GGAC	
	Intra	Inter	Intra	Inter
$\Delta\Delta E_{LJ}$	−1.0 (−5.0)	11.9 (3.6)	−1.8 (−9.2)	13.8 (6.1)
$\Delta\Delta E_{Coul}$	−61.2 (−0.2)	−47.2 (7.8)	−23.2 (6.4)	−82.3 (−10.5)

Calculated intra- and interstrand van der Waals ($\Delta\Delta E_{LJ}$) and Coulomb ($\Delta\Delta E_{Coul}$) interaction energy differences between lowest energy ftf and sheared forms are given in kcal mol^{-1}. The numbers in parenthesis correspond to base-base interactions only.

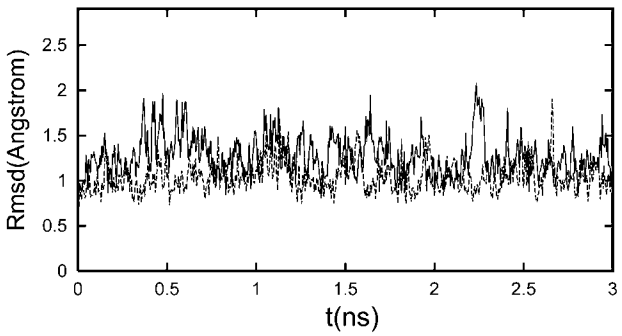

FIGURE 4 The RMSD time course of the 3-ns data gathering time interval with respect to the start structure for the (rGCG**GA**CGC)$_2$ sequence with the tandem GA mismatch in the ftf arrangement (*black line*), and the (rGGC**GA**GCC)$_2$ sequence in sheared conformation (*dashed line*).

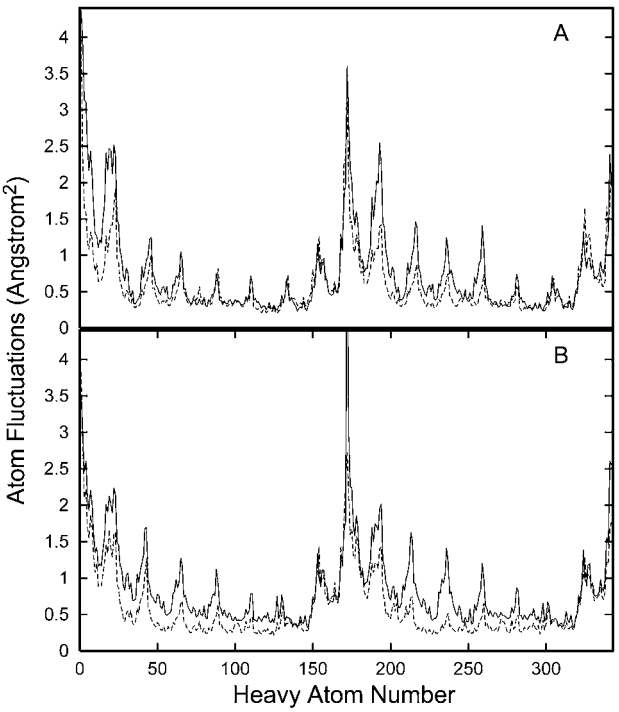

FIGURE 5 Average heavy atom position fluctuations during molecular dynamics simulations (2- to 4-ns data gathering time interval) of (*A*) (rGCG**GA**CGC)$_2$ and (*B*) (rGGC**GA**GCC)$_2$ RNA molecules, respectively (fluctuations of ftf-form, *continuous line*; sheared form, *dashed line*).

than experimental estimates of 2–3 kcal mol^{-1}. Interestingly, the calculations that assume an internal dielectric constant, $\varepsilon_{in} = 2$, are closer to experiment than the standard, $\varepsilon_{in} = 1$, calculations. Similar to the results on comparing minimized lowest energy substates, the trajectory analysis indicates the sheared form overall to be favored by van der Waals and polar and nonpolar solvation contributions, but disfavored by Coulomb interactions. The trajectory analysis predicts that differences in the total electrostatic contributions make a significant contribution to the sequence context-dependence of the relative stability of sheared vs. ftf tG:A mismatches. In the case of the C**GA**G context, electrostatic contributions overall slightly favor the sheared form, whereas in the case of the G**GA**C context, the ftf form is electrostatically favored over the sheared form.

DISCUSSION

The analysis of the energetic origins that determine conformational preferences of structural motifs in nucleic acids is essential to better understand their function in biological RNA-containing molecules. Preferably, computational studies on nucleic acids should include surrounding solvent and ions explicitly. However, explicit solvent simulations are currently limited to small timescales (nanoseconds). Calculations on free energy differences for two conformational substates of a biomolecule are, in principle, possible using thermodynamic integration or perturbation methods. Such approaches, however, require the determination of a thermodynamic average over solute and all solvent degrees of freedom for which convergence is difficult to achieve. Continuum solvent methods, although principally less accurate than explicit solvent simulations, allow the calculation of an average solvent polarization for a given molecule conformation and, in turn, an estimate of the degree of stabilization by the surrounding solvent and ion atmosphere. Such calculations can give a hint as to what interactions stabilize or destabilize a particular structural topology of a given motif.

As a basis for the present calculations, two NMR-derived RNA structures with identical base composition, but slightly different sequence, and the tG:A motif either in the sheared form (CGAG context, Santa-Lucia and Turner, 1993) or the ftf form (GGAC context, Wu and Turner, 1996), were chosen. This pair of structures is a very useful reference and basis for the present theoretical study, since both structures have been determined under similar conditions, using the same experimental method, and are composed of the same nucleotides. Other influences besides of the sequence context that may affect the relative stability of tG:A mismatch conformations such as solution or crystallization conditions and contacts to other RNA elements in folded RNAs can be excluded.

TABLE 4 Continuum solvent analysis of MD trajectories

Sequence time interval	(rGGC**GA**GCC)$_2$, 3 ns	(rGCG**GA**CGC)$_2$, 3 ns
$\Delta E_{totsalt}$ (ftf vs. sheared form)	10.7 ± 3 (8.0 ± 2)	−13.2 ± 7 (−4.2 ± 2)
ΔE_{tot} (without salt)	11.1 ± 3 (8.4 ± 2)	−12.8 ± 7 (−3.5 ± 2)
ΔE_{Estat}	3.6 (1.0)	−19.2 (−10.0)
ΔE_{Coul}	−91.3 (−45.7)	−60.7 (−30.4)
$\Delta E_{rePB(+0.15\ M\ salt)}$	94.8 (46.5)	41.5 (20.1)
ΔE_{LJ}	8.3	5.6
ΔE_{SASA}	0.5	0.5
ΔE_{TA}	−1.2	0.8

Energetic contributions are given in kcal mol^{-1} and correspond to the difference between the trajectory analysis of the face-to-face vs. sheared tandem G:A topologies with the same sequence. Numbers in parenthesis correspond to calculations using an internal dielectric constant $\varepsilon_{in} = 2$ (otherwise $\varepsilon_{in} = 1$). Errors have been calculated by comparing the total energy differences obtained for the first part of the trajectory with differences obtained for the second part of the trajectories. ΔE_{TA} includes all bonded energy contributions (bond length, bond angle, and torsion-angle energy contributions). The reaction field energies, $\Delta E_{rePB(+0.15\ M\ salt)}$, were calculated including 0.15 M monovalent salt and solving the nonlinear FDPB (Sharp and Honig, 1990).

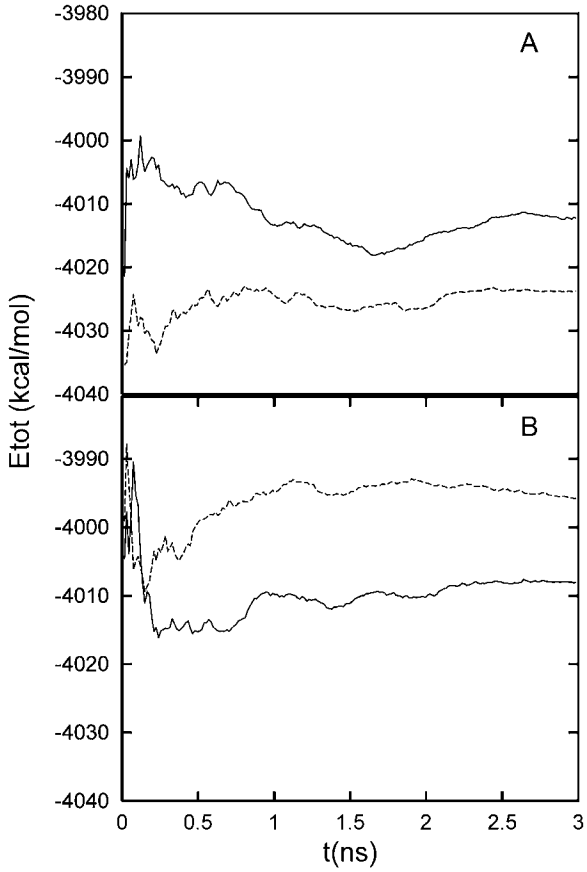

FIGURE 6 Cumulative average of calculated mean total energies from the MM/PBSA-trajectory analysis (final averages over 200 conformers). (*A*) Simulation of the (rGGC**GA**GCC)$_2$ structures; (*B*) simulation of the (rGCG**GA**CGC)$_2$ structures. The continuous lines indicates the results for the ftf-forms (*dashed line*, sheared form).

In the present study two different continuum solvent analysis methods have been applied to study the conformational preference of tandem guanine:adenine mismatches in RNA. In the systematic conformational analysis approach, energy-minimized substates for the motif have been generated by applying constraint energy minimization to systematically generate various backbone torsion-angle combinations defined by a set of window constraints followed by free minimization to reach the closest energy minimum substate. In the case of the MM/PBSA approach (Srinivasan et al., 1998; Kollman et al., 2000), ensembles of tG:A mismatch conformations have been generated using molecular dynamics simulations including ions and water explicitly followed by postprocessing the trajectories using the same force field and continuum solvent method as for the analysis of substates. Both approaches complement each other, in that the substates analysis method samples systematically a wider range of conformers, whereas the MM/PBSA approach yields more or less randomly sampled conformers close to one or a few substates. It needs to be stressed that a perfect agreement between substate analysis and trajectory analysis is not expected. In the former (and faster) approach, a greater variety of conformers can be sampled, but these are only represented as energy minima. In contrast, during molecular dynamics simulations, not only energy minima but conformers compatible with a thermodynamic ensemble in the presence of explicit solvent have been generated. However, even at the nanosecond timescale the simulated structures stay relatively close to one substate which is given by the starting structure. In the present study, the lowest energy substates that showed best agreement with available experimental structures were chosen as start conformations. An additional difference between the conformational searches using JUMNA (Lavery et al., 1995) and the analysis of MD trajectories is that all bond lengths (and valence angles within nucleobases) are kept at their optimal values in the JUMNA approach (Lavery et al., 1995). The reduced set of conformational variables allows a fast convergence of the energy-minimization calculations. In the MD simulations, bond lengths between heavy atoms were allowed to vary, and, in particular, atoms that belong to nucleobases can undergo more fluctuations.

Despite these differences, both approaches yielded qualitatively similar results on the energetic contributions that stabilize or destabilize sheared and face-to-face topologies of tG:A mismatches and on the experimentally observed sequence context effect. The analysis of energy-minimized substates yielded a more significant stabilization of the sheared form in the C**GA**G context but still predicted a lower energy for the sheared vs. ftf forms in the case of the G**GA**C context. However, the difference was smaller by ~3 kcal mol^{-1} for the latter case. The MD simulations of the **tG:A** motif in the two sequence contexts showed an overall greater flexibility of the face-to-face topology that indicated an entropic favorization of the face-to-face topology not accounted for in the energetic evaluation of the substates. Accounting for such a contribution would overall shift the stability difference between the two topologies in favor of the face-to-face form in better agreement with the experimental observation. The MM/PBSA trajectory analysis yielded stability differences between the two tandem G:A forms that have a sign in agreement with experiment but are larger than experimental estimates (Wu and Turner, 1996). Interestingly, we found that using a larger internal dielectric constant of $\varepsilon_{in} = 2$ still yields the correct sign, but also somewhat smaller energy differences between the two mismatch topologies in the two sequence contexts that are probably more realistic. It should be noted that the force field of Cornell et al. (1995) has been designed for simulation studies using a dielectric constant $\varepsilon_{in} = 1$. Since for the present analysis method the explicit solvent from the simulation has been replaced by a continuum without reducing or changing the degrees of freedom of the solute molecule, the use of $\varepsilon_{in} = 1$ for the interior of the molecule is a consistent choice. However, it has been found in other applications of the MM/PBSA method, for example Wang

et al. (2001), that a trajectory analysis with a larger dielectric constant often yields results in better agreement with experiment than with $\varepsilon_{in} = 1$.

The qualitative agreement between the conformational search approach and the MM/PBSA method allows to identify trends concerning the energetic origins for the stabilization of the two topologies. Both approaches predict a stabilization of the sheared topology due to van der Waals and nonpolar solvation contributions. This is to some degree expected, considering the more compact shape of the sheared tG:A arrangement. The MM/PBSA results, as well as comparison of lowest energy sheared and ftf substates, indicated that Coulomb interactions generally favor the face-to-face pairing, although the number of polar contacts such as hydrogen bonds is similar in both topologies (depending on the substate, even more hydrogen bonds are possible in the sheared topology). However, the less compact shape of the face-to-face topology (with an on average greater distance between opposite strands) results in an overall smaller Coulomb repulsion of the phosphate groups on opposite strands. In the sheared form, this is, in part, compensated by a more negative reaction field energy which stabilizes the sheared form. These trends are seen for both sequence contexts. A role of electrostatic contributions to stacking for the context effect has been suspected already by Wu and Turner (1996). For the lowest energy ftf and sheared conformations, electrostatic and van der Waals interactions were further split into inter- and intrastrand contributions. The calculated intra- and interstrand contributions to the van der Waals interactions of the RNA molecules were similar in both sequence contexts. The calculated van der Waals favorization of the sheared topology was found to be mostly due to interstrand interactions. The analysis showed further that the Coulomb part of the interstrand base-base interaction significantly favors the sheared arrangement in the C**GA**G context, whereas the same interaction favors the ftf arrangement in the G**GA**C context. This result is in line with the results of the MM/PBSA trajectory analysis that predicted an overall electrostatic favorization of the sheared form in the C**GA**G context and ftf form in the G**GA**C context. It indicates that electrostatic interactions make the most significant contribution to the sequence context effect on the conformational preference of tandem G:A mismatches. The two continuum solvent approaches applied to tandem G:A mismatches could also be useful to analyze and better understand conformational preferences of other nucleic acid motifs. It might also be possible to use the methodology to make predictions on sequence context effects and how conformational preferences might change upon changing the environment of the nucleic acid for, example, in complexes with proteins or organic ligands.

We acknowledge helpful discussions with F. Pineda.

This work was supported, in part, by a grant from the Deutsche Forschungsgemeinschaft (ZA 153/3-1) to M.Z.

REFERENCES

Auffinger, P., and E. Westhof. 1997. Rules governing the orientation of the 2′-hydroxyl group in RNA. *J. Mol. Biol.* 274:54–63.

Ban, N., P. Nissen, J. Hansen, P. B. Moore, and T. A. Steitz. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science.* 289:905–920.

Bondi, A. 1964. Van der Waals volumes and radii. *J. Phys. Chem.* 64:441–451.

Cate, J. H., A. R. Gooding, E. Podell, K. Zhou, B. L. Golden, C. E. Kundrot, T. R. Cech, and J. A. Doudna. 1996a. Crystal structure of a group I ribozymes domain: principles of RNA packing. *Science.* 273:1678–1685.

Cate, J. H., A. R. Gooding, E. Podell, K. Zhou, B. L. Golden, A. A. Szewczak, C. E. Kundrot, T. R. Cech, and J. A. Doudna. 1996b. RNA tertiary structure mediation by adenosine platforms. *Science.* 273:1696–1699.

Cornell, W. D., P. Cieplak, C. I. Bayley, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1995. A second generation force field for simulation of proteins, nucleic acids and organic molecules. *J. Am. Chem. Soc.* 117:5179–5197.

Gautheret, D., D. Konings, and R. R. Gutell. 1994. A major family of motifs involving G:A mismatches in ribosomal RNA. *J. Mol. Biol.* 242:1–8.

Gutell, R. R., N. Larsen, and C. R. Woese. 1994. Lessons from an evolving RNA: 16S and 23S rRNA structures from a comparative perspective. *Microb. Rev.* 58:10–26.

Guzman, R. N., R. T. Turner, and M. F. Summers. 1998. Protein-RNA recognition. *Biopolymers.* 48:181–195.

Hawkins, G. D., C. J. Cramer, and D. G. Truhlar. 1995. Pairwise solute descreening of solute charges from a dielectric continuum. *Chem. Phys. Lett.* 246:122–129.

Hawkins, G. D., C. J. Cramer, and D. G. Truhlar. 1996. Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. *J. Phys. Chem.* 100:19824–19839.

Heus, H. A., S. S. Wijmenga, H. Hoppe, and C. W. Hilbers. 1997. The detailed structure of tandem G:A mismatched basepair motifs in RNA duplexes is context-dependent. *J. Mol. Biol.* 271:147–158.

Jayaram, B., D. Sprous, and D. L. Beveridge. 1998. Solvation free energies of biomacromolecules: parameters for a modified generalized Born model consistent with the Amber force field. *J. Phys. Chem.* 102:9571–9576.

Jorgensen, W. L., J. Chandrasekhar, J. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

Kollman, P. A., I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, and T. E. Cheatham 3rd. 2000. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* 33:889–897.

Lavery, R., K. Zakrzewska, and H. Sklenar. 1995. JUMNA (junction minimization of nucleic acids). *Comput. Phys. Com.* 91:135–158.

Madura, J. D., M. E. Davis, R. Wade, B. A. Luty, A. Ilin, A. Anosiewicz, M. K. Gilson, B. Bagheri, L. Ridgway-Scot, and J. A. McCammon. 1995. Electrostatics and diffusion of molecules in solution: simulations with University of Houston Brownian dynamics program. *Comput. Phys. Commun.* 91:57–95.

Pearlman, D. A., D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, S. Debolt, D. Ferguson, G. Seibel, and P. A. Kollman. 1995. AMBER; a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.* 91:1–41.

Pley, H. W., D. S. Lindes, and D. B. McKay. 1994. Three-dimensional structure of a hammerhead ribozyme. *Nature.* 372:68–74.

SantaLucia, J., and D. H. Turner. 1993. Structure of $(rGGCGAGCC)_2$ in solution from NMR and restraint molecular dynamics. *Biochemistry*. 32:12612–12623.

Schlünzen, F., A. Tocilj, R. Zarivach, J. Harms, M. Gluehmann, D. Janell, A. Bashan, A. Bartels, I. Agmon, F. Franceschi, and A. Yonath. 2000. Structure of functionally activated small ribosomal subunit at 3.3 Å resolution. *Cell*. 102:615–623.

Schlünzen, F., R. Zarivach, J. Harms, A. Bashan, A. Tocilj, R. Albrecht, A. Yonath, and F. Franceschi. 2001. Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature*. 413:814–821.

Sharp, K. A., and B. Honig. 1990. Calculating total electrostatic energies with the nonlinear Poisson-Boltzmann equation. *J. Phys. Chem*. 94: 7684–7692.

Shrake, A., and J. A. Rupley. 1973. Environment and exposure to solvent of protein atoms: lysozyme and insulin. *J. Mol. Biol*. 79:351–365.

Sitkoff, D., K. A. Sharp, and B. Honig. 1994. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem*. 98:1978–1988.

Srinivasan, J., J. Miller, P. A. Kollman, and D. A. Case. 1998. Continuum solvent studies of the stability of RNA hairpin loops and helices. *J. Biomol. Struct. Dyn*. 16:671–682.

Srinivasan, J., M. W. Trevathan, P. Beroza, and D. A. Case. 1999. Application of a pairwise generalized Born model to proteins and nucleic acids: inclusion of salt effects. *Theor. Chem. Acc*. 101:426–434.

Still, W. C., A. Tempczyk, R. C. Hawley, and T. Hendrikson. 1990. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc*. 112:6127–6129.

Tsui, V., and D. A. Case. 2000. Molecular dynamics simulations of nucleic acids with a generalized Born solvation model. *J. Am. Chem. Soc*. 122:2489–2498.

Wang, W., W. A. Lim, A. Jakalian, J. Wang, J. Wang, R. Luo, C. Bayly, and P. A. Kollman. 2001. An analysis of the interactions between the Sem-5 SH3 domain and its ligands using molecular dynamics, free energy calculations, and sequence analysis. *J. Am. Chem. Soc*. 123:3986–3994.

Williams, D. J., and K. B. Hall. 1999. Unrestrained stochastic dynamics simulations of the UUCG tetraloop using an implicit solvation model. *Biophys. J*. 76:3192–3205.

Williams, D. J., and K. B. Hall. 2000a. Experimental and theoretical studies of the effects of deoxyribose substitutions on the stability of the UUCG tetraloop. *J. Mol. Biol*. 297:251–265.

Williams, D. J., and K. B. Hall. 2000b. Experimental and computational studies of the G[UUCG]C RNA tetraloop. *J. Mol. Biol*. 297:1045–1061.

Wimberley, B. T., D. E. Brodersen, W. M. J. Clemons, R. J. Morgan-Warren, A. P. Carter, C. Vonrhehein, T. Hartsch, and V. Ramakrishnan. 2000. Structure of the 30S ribosomal subunit. *Nature*. 407:327–339.

Wu, M., J. SantaLucia, and D. H. Turner. 1997. Solution structure of $(rGGCAGGCC)_2$ by two-dimensional NMR and iterative relaxation matrix approach. *Biochemistry*. 36:4449–4460.

Wu, M., and D. H. Turner. 1996. Solution structure of $(rGCGACGC)_2$ by two-dimensional NMR and iterative relaxation matrix approach. *Biochemistry*. 35:9677–9689.

Zacharias, M., and H. Sklenar. 1999. Conformational analysis of single-base bulges in A-form DNA and RNA using a hierarchical approach and energetic evaluation with a continuum solvent model. *J. Mol. Biol*. 289:261–275.

Zacharias, M. 2001. Conformational analysis of DNA-trinucleotide-hairpin-loop structures using a continuum solvent model. *Biophys. J*. 80:2350–2363.

Zhang, X., B. L. Gaffney, and R. A. Jones. 1998. 15_N NMR of RNA fragments containing specifically labelled tandem G:A pairs. *J. Am. Chem. Soc*. 120:6625–6626.